

EUCALL

The European Cluster of Advanced Laser Light Sources

Grant Agreement number: 654220

Work Package 5 – UFDAC

Deliverable D5.3

Report on online processing of digitizer data

Lead Beneficiary: European XFEL

Authors: Patrick Geßler, Bruno Fernandes, Hamed Sotoudi Namin,
Gabor Kavai, Kwinten Nelissen, Florin Negoita, Mihai Cuciuc, Stanimir Kisyoov, Soiciro Aogaki

Due date: 30.09.2018

Date of delivery: 29.09.2018

Project webpage: www.eucall.eu

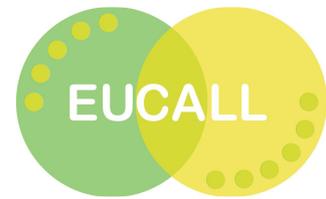
<i>Deliverable Type</i>	
R = Report DEM = Demonstrator, pilot, prototype, plan designs DEC = Websites, patents filing, press & media actions, videos, etc. OTHER = Software, technical diagram, etc.	R
<i>Dissemination Level</i>	
PU = Public, fully open, e.g. web CO = Confidential, restricted under conditions set out in Model Grant Agreement CI = Classified, information as referred to in Commission Decision 2001/844/EC	PU



LUND UNIVERSITY



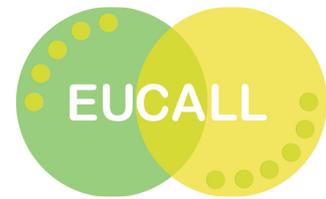
This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654220



Contents

1. Introduction.....	3
2. Overview of applications and challenges.....	3
3. Digitizer and similar technologies and processing capabilities.....	4
3.1 Overview of ADC and Digitizers	4
3.2 Sensors	6
4. Timing and synchronization	8
4.1. White Rabbit	8
4.2 Distributed triggering over White Rabbit	8
4.3 Network Topology.....	9
4.4 Main results	9
4.5 Conclusion and outlook	10
5. Identified demand for low-latency online processing functionality.....	11
5.1 Real-time distributed feedback systems	11
5.2 VETO system	12
5.3 Data reduction	14
6. Implemented processing algorithms.....	14
6.1 Energy detection	14
6.2 Peak detection	17
6.3 Zero suppression.....	18
6.4 Pile-up compensation	19
7. Usage of developed solutions	26
7.1 Measurements at PETRA III and European XFEL.....	26
7.2 Measurements from ELI-NP	29
8. Conclusion	29
References and publications	30





1. Introduction

The repetition rates and data volumes of modern light sources cause increasing challenges for the development of detectors, diagnostics tools and equipment. For an efficient usage of beamtime at the optical laser and FEL RIs, high performance data acquisition (DAQ) systems must provide processing capabilities close to the equipment, enabling data analysis and validation at acquisition time. The complexity of integrating and correlate data from the different sources requires specialize hardware and firmware development, ranging from on-ASIC memory, fast digitizers and high throughput DAQ hardware and software.

In this report, a summary of integrated digitizer platforms and implemented FPGA algorithms for online processing of data is presented. Contributions from European XFEL, ELI and ESRF are included.

2. Overview of applications and challenges

Many detectors provide transient signals like pulses and spectral data (e.g. Time of Flight - ToF). Those detectors are connected to digitizers implementing analogue to digital converters (ADCs). FPGA processing would allow to calculate the relevant parameters of the signals in real-time with low latency. This would provide multiple advantages and possibilities:

- The amount of data would be significantly reduced (especially for high-speed digitizers > 1 GHz). Furthermore, the real-time and low latency data could directly be used for
- feedback loops to optimize laser performance or stability as well as
- for VETO systems, which would allow to select the most promising data sets captured (especially if the detector storage pipeline or transfer bandwidth was limited).

When taking into account the large data volumes delivered by the detectors (of the order of 10 to 15 Gb per second per detector) as well as the short interval between runs, the task of performing data processing at the front-end and communicate this information between the different hardware elements that are involved in the DAQ system is no small feature.



3. Digitizer and similar technologies and processing capabilities

3.1 Overview of ADC and Digitizers

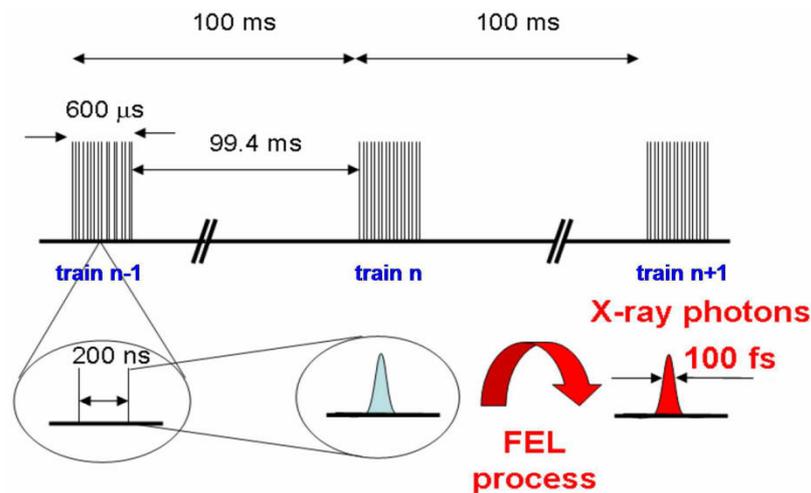


Figure 1: Structure of the European XFEL pulse train

The European X-ray Free Electron Laser facility will generate intense ultra-short coherent x-ray flashes paced by 220 ns and each with a width of less than 100 femtoseconds. They are grouped into trains of up to 2700 pulses within 600 μs with a 10 Hz repetition rate. Bandwidths of 10 GBytes of data per second are expected from 2D pixel detectors, like AGIPD and LPD, while other detector types, like systems based on fast digitizing analog-to-digital converters, can go up to 60 MBytes per channel. In this section we will focus in these later detectors, here on refer to as digitizers.

The main Hardware platform is based on MicroTCA.4. It offers high bandwidth communication between boards and CPU via PCI Express (PCIe) and between boards via point-to-point communications. It also offers dedicated High-speed data throughput and processing units, necessary to collected and reduces the amount of data to be store, is available in the form of CPUs, DSP and Boards/Digitizers with FPGA. The platform offers configurable low jitter clocks lines distribution together with an eight bussed M-LVDS lines for less demanding and triggers signals in the backplane (see Figure 2).

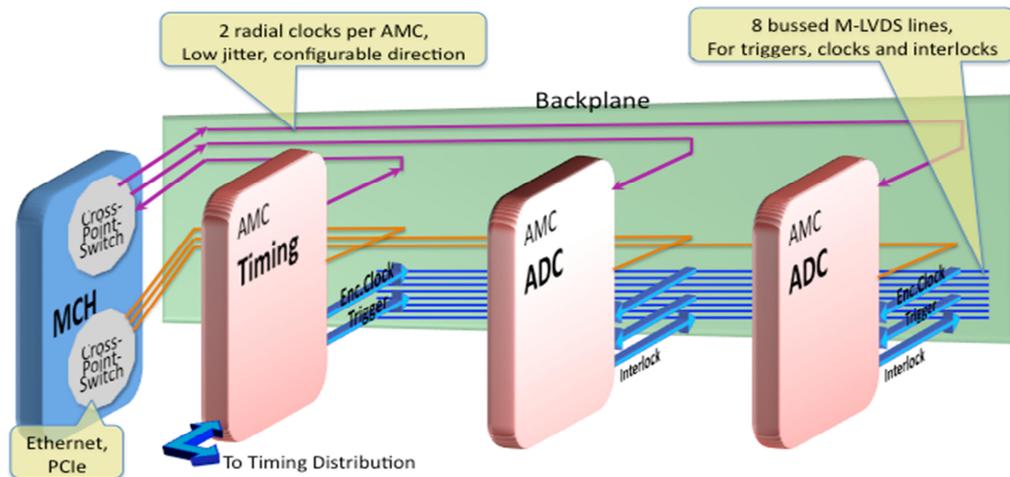


Figure 2: Overview of MicroTCA communication channels

The digitizers in European XFEL vary in resolution and sampling frequency (see Figure 3). All allow for raw data collection and algorithms for data processing are also available and are described in section 5. The following is an overview of the available digitizers:

- **SIS8300** board is a ten channel 125 MS/s digitizer with a 16 bit resolution, design according to the μ TCA for Physics standard. The board has a Virtex-5 Xilinx FPGA which implements a 4 lane PCI Express interface. Other features include a 4Gb DDR2 Memory, 2 DAC channels, external clock and trigger sources and dual SFP card cage for Multi Gigabit Link communication
- For more high frequency signals, **ADQ412** and **ADQ7** boards from **SP Devices** are available. The former provides 4 channels, with 2GHz sampling frequency and 12 bit resolution; the device can be configured to use only 2 channels and sample at a 4 GHz rate. Similarly, the ADQ 7 is a two channel 5GHz with 14 bit resolution board that can be configured to work at 10 GHz sampling frequency using only one channel.

Both product groups are not only used at the European XFEL, but also at different facilities related to light sources and beyond. Therefore developed processing algorithms could easily be used also at those facilities.

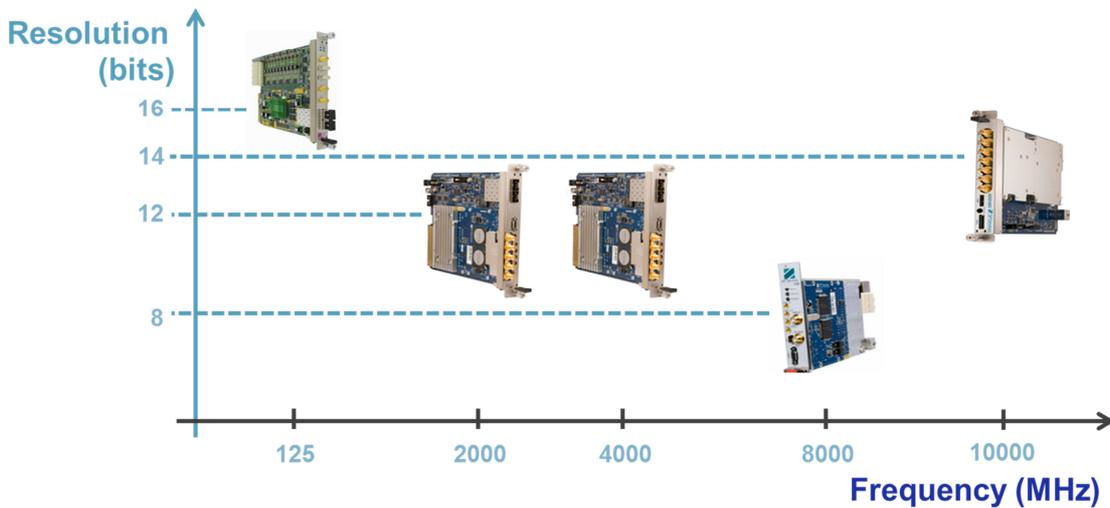


Figure 3: Digitizers in use at European XFEL

3.2 Sensors

In experiments using light at various wavelengths, most of the diagnostics are based on CCDs that provide 2D images. They are used mostly to measure the scattered/produced radiation following the beam interaction with target/sample and thus to infer the properties of the later or of the interaction itself. Convertors, screens, image intensifiers etc. are place typically in front of CCDs in order to transform the measured radiation in the range of sensitivity of the employed CCDs. Additionally, CCDs are used in measuring various parameters of the main or of the probe/auxiliary beams: (spatial dimensions/profile, pointing, spectrum, etc.) in order to ensure stability and meaningful data interpretation. Depending mainly on incident beam repetition rates, very large amounts of digital data is generated by CCDs.

For the measurement of the energy spectrum of hard x- and gamma-rays, semiconductor detectors are most suitable. Silicon or germanium crystals, usually cooled to liquid nitrogen temperatures, followed by charge preamplifiers with first stage also cooled, provide a signal with amplitude proportional to deposited energy. Signals are feeding fast digitizers (100 - 1000 MS/s) with high resolution (14 bits) that produce 1D data that can be stores or further processed in FPGA/GPU for reduction before storage. Small size semiconductor detectors, operated at room temperatures (as shown in the right hand side of the figure below) are widely used for charge particle detection and energy measurement.



Figure 4: Example of detectors

Scintillators are based on light emission following the energy deposition due a particle interaction. Light (a rather limited number of photons) is emitted and transform in electric signal by a photomultiplier (PMT) feeding a fast (~ 1 GS/s), high resolution ($\sim 12-14$ bits) digitizer such that the deposited energy is obtained from signal integration. Implementation of integration and time pick-up algorithms is available commercially with rejection of events that are overlapping (pile-up events). The capability to process also pile-up events in order to obtain (even with degraded resolution) the energy and arrival time is beneficial in many cases, such as when total rate is high while the number of events of interested is low or when events come in burst. In high power laser experiments such conditions may occur in vicinity of target due to nuclear activation of short lived states. Commercial scintillators coupled to photomultipliers have been proved able to operate, despite strong x-ray flash and electromagnetic pulse, inside an air bubble as shown in Figure 5 below.

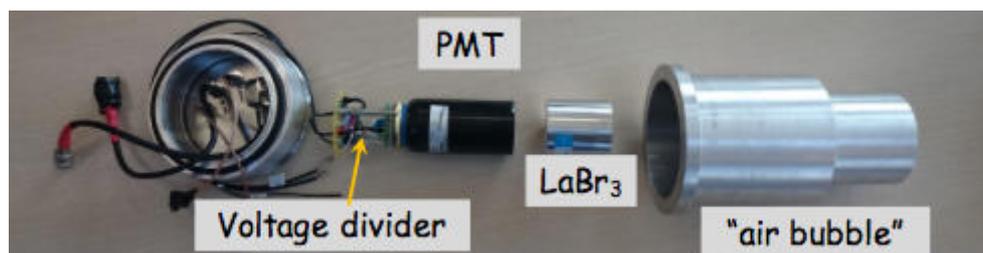
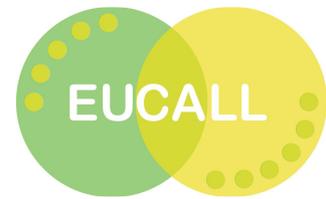


Figure 5: Scintillators couple to photomultipliers

3.2.1 Integrating current transformers

In the laser electron acceleration experiments, energetic electron bunches will be produced from laser plasma interactions in various gas targets. The measurement of the total charge of the accelerated electron bunch is an important diagnostic that has to be performed in order to understand and control the interaction mechanisms and correlate experiments with simulations. Values in the range of pC to tens of nC are expected to be obtained in experiments and integrating current transformers (ICT) will be used as part of the diagnostic equipment usually shortly after the gas target, before more sophisticated diagnostic such as electron spectrometers. The ICTs (see figure below) are non-interceptive diagnostics that have to be coupled with appropriate electronics and further with digitizers.



4. Timing and synchronization

In the application of high-speed digitization and real-time feedback systems, timing and synchronization plays a crucial role. Here usually three aspects are of high importance: (1) provide a distributed absolute time reference (e.g. time stamps), (2) provide a common reference clock (e.g. frequency), and (3) to provide triggers (like start actions to synchronize acquisition tasks). While different solutions of such timing systems exist, the White Rabbit timing system is a relative recent solution, which has to potential of being adopted in a larger number of laser light source facilities. However, the system does not provide triggers natively. Therefore further investigations, developments and tests had been carried out within the scope of the UFDAC work and should be presented briefly in this section.

4.1. White Rabbit

Diagnostic devices, actuators, etc. have to be timed against the arrival of the laser pulse with high accuracy. For this it is necessary to ensure the reliable and precise distribution of electric trigger signals – produced by laser sources –over the facility. The devices which need accurate timing should synchronize their operation to these trigger pulses. A possible instrument for this purpose is the White Rabbit system developed by the engineers of CERN and other research institutes. The White Rabbit timing system is able to synchronize the time of its nodes with sub-ns accuracy to a reference clock (for e.g. GPS). On the other hand White Rabbit is not designed to redistribute trigger signals from an arbitrary source towards an arbitrary destination. However the White Rabbit optical network is capable to distribute conventional Ethernet packets, besides the packets responsible for time synchronization. Thus the network is capable of dealing with control and data packets, which capability we will explore for setting up a distributed triggering system over White Rabbit.

4.2 Distributed triggering over White Rabbit

The aim of this work within UFDAC was to acquire electrical trigger signals from an input source, e.g. a laser system, distribute them over the White Rabbit optical network and reproduce them for an external device. In the course of this work to topology as presented in Figure 6 is implemented.

At the beginning of the development is set up the cross compiler, allowing the faster code compilation in PC. The coding was performed in C and compiled with a cross-compiler targeting the ZEN ARM platform.



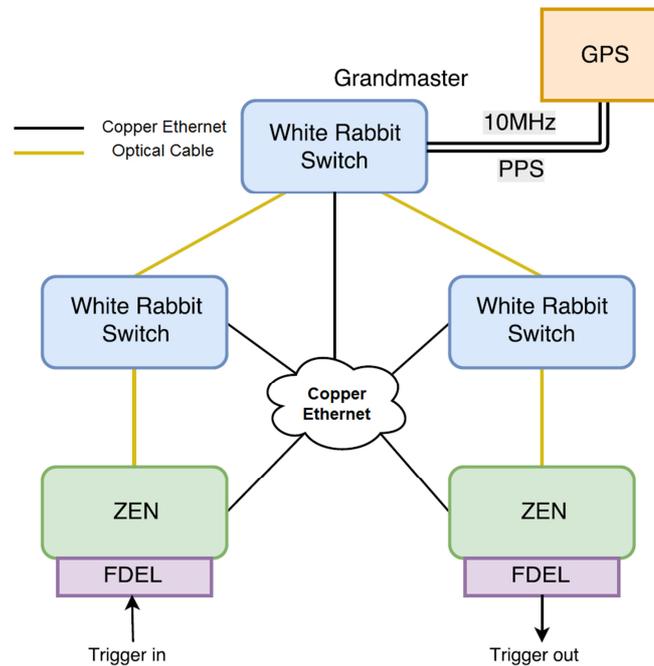


Figure 6: The topology of the measurement

4.3 Network Topology

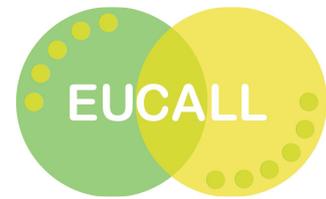
The acquiring and re-generation of pulses was made on the Seven Solution’s White Rabbit ZEN with an FMC Fine-Delay card. Between the two nodes the connection was realized with the White Rabbit optical network and with a conventional copper Ethernet network. While the optical network is responsible for synchronization to the reference clock, the copper wired Ethernet with TCP socket connection was used to distribute the timestamps of trigger pulses.

4.4 Main results

During the jitter measurements the frequency and phase difference between input and the reconstructed trigger pulses are measured in the long term period (thousands of signal sequences). The input trigger’s period was constantly 30 ms during measurement. The jitter measurement result was 37.343ps RMS which was slightly more what we expected based on the FMC-Fine-delay card’s specification: 30 ps RMS.

During the performance measurements the number of missed trigger pulses is counted over a period of 5-20. The maximum achieved frequency was 500 Hz, where the dropped trigger pulse number is 0,038% and the timed out trigger pulse number is also 0,038%. At lower repetition rates no pulses are missed.

Using of the other output channels of the FMC Fine Delay card in parallel is not affected the measurement.



4.5 Conclusion and outlook

The above-mentioned measurements showed that the trigger distribution implemented in software has limitation with respect to the achievable frequency. This problem can be resolved by implanting the trigger distribution in hardware. However, the White Rabbit ZEN firmware is not a part of the White Rabbit open source project disabling hardware development from customer side. With modification of the VHDL code in FPGA, the maximum frequency can potentially be increased significantly.

The main reasons for the low maximum frequency are the next latencies:

- Copying the timestamps from the FMC Fine Delay card's circular buffer
- Adding operation with the timestamps
- TCP/IP stack handling by the operating system
- FMC Fine Delay card's output setup dead time

In order to enlarge the maximum trigger frequency, it would be great testing this project with the White Rabbit NIC driver, i.e. using the optical network for data communication. Meanwhile on the ZEN we successfully set-up the NIC for the WR optical network, new measurements are being planned. However, the results are expected to be very similar to the one obtained from copper Ethernet trigger distribution, because of the above-mentioned reasons of latencies. Currently we are implementing distributed triggering on Seven Solution's SPEC cards, which are PCIe extension cards which are fully open source.

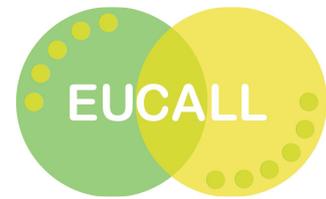
Reaching the higher trigger frequencies, we recommend using the CERN's:

- Trigger Distribution project
- RF Distribution project

The issue with those projects is that they are in demo phase and the hardware elements are not in commercial production. The Trigger Distribution is realized with SVEC + FMC Fine Delay Card, the timestamp and the Ethernet packet handling is solved by the deterministic Mock Turtle FPGA core. With this optimization, CERN reached an 89 μ s trigger distribution latency. In the CERN's RF Distribution project the SVEC + DDS600M FMC card solves the frequency updating without any glitches.

Porting the CERN's Trigger Distribution project to the SPEC card would be interesting, because in that case, it wouldn't require extra hardware elements.





5. Identified demand for low-latency online processing functionality

5.1 Real-time distributed feedback systems

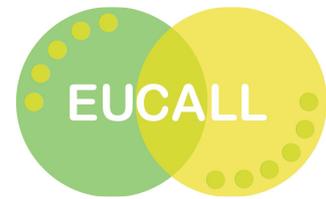
Preserving the parameters of the laser system is a key element in laser – driven experiments as the beam parameters (e.g. pointing drift, contrast, energy, intensity distribution, pulse duration, spectrum, etc) variation strongly influence the light - matter interaction. Typically, up to 5 hours are needed daily in order to warm up the laser system and align the system prior to start performing the experiments. Automatic alignment and protection procedures are envisaged in order to track and align the laser beams in the high-power laser system and determine on-line potential dangerous situations (e.g. abnormal beam intensity profiles) that can irreversibly damage expensive components of the laser system.

In order to allow the safe operation and to minimize the warm up time of the laser equipment, the laser beam properties have to be precisely monitored. Further on, piezo driven mirrors have to be actuated in automatic loops in order to preserve the desired beam path and machine safety measures triggered to protect the laser as soon as a potential danger is encountered. Coupled with advanced algorithms based on neural networks, genetic or heuristic algorithms, feedback loops can be implemented to drive various parameters in the laser system (e.g. beam pointing, spectrum, chirp, compression) or of the target (position relative to focus) in order to stabilize or to improve the final result (e.g. energy spectrum of generated radiation).

In general, the implementation of the data evaluation and feedback loop does not need a real-time response as a new laser shot can be set on demand, when the next parameters have been calculated and updated into the laser system. However, if a primary data has to be collected from various sensors and large numbers of iterations are needed in order to converge to optimized parameters, the implementation of the algorithms and feedback loops may need real-time response in order to benefit from the maximum repetition rate of the laser system. As the laser front-ends operate typically at 10 Hz, the detection of the beam parameters (e.g. position, intensity profile) together with the control (e.g. mirrors position, safety measure trigger) has to be performed at the same repetition rate with high priority and real-time response. Because of the real-time response needed, with high priority and due to the requirement to implement advanced algorithms than need to determine the control action at fast repetition rate, the implementation will have to be performed in an FPGA.

A method under development at ELI-NP is using scintillators for gamma spectroscopy of short-lived nuclear isomers as diagnostics for laser accelerated particles such as proton or gamma flux emitted with energy above a certain threshold in the solid angle covered by a secondary target viewed scintillator. Thus, processing on-line the digitized data the produced





particle flux can be monitored and kept at desired level acting on laser parameters (e.g. wavefront, energy) or target parameters (such as pressure or time to open the valve in case of a near-critical gas target. After high power laser shot and scintillator signal recovery, still a large number of gamma rays are hitting the detectors during a short time (up to milliseconds) when short-lived nuclear isomers are produced.

5.2 VETO system

In the example of European XFEL, different types of detectors and diagnostic sensors will be used at the experiments and along the beam lines. For those systems providing information for each individual bunch, a maximum of 2700 data sets per 10 Hz are expected. Especially of detectors providing 1D or even 2D data sets per bunch will produce a high amount of data bandwidth, which has to be transported through the individual data acquisition (DAQ) chains till it will be stored on a file system.

For some detectors the number of data sets is limited. That implies, that it is not possible to record all possible 2700 bunches. However, currently all those detectors provide a way to remove unwanted data sets in an early stage in the detector head and therefore frees up space for other, possibly better, data sets. In this way a much more efficient use of the limited storage could be achieved.

This described mechanism is called VETO, as a low latency signal providing a decision, if one data set should be removed from the storage acts as a veto right to the detector. For those detector or diagnostic systems, which are able to provide information for all possible bunches, the VETO system could be used to reduce the amount of data to be transmitted through the DAQ chain in an early stage, if the data sets which are rejected are not worth to look at or to be saved anyway.

An overview of the VETO system interfacing elements is shown in Figure 7. The interfacing elements include:

- **VETO Sources:** devices that deliver information relevant in order to make a decision if a data set of a connected VETO user could be rejected or removed from the storage. Examples of possible VETO sources are ADC and digitizer based detectors, as described in the report. They are able to provide important information with very low latency.
- **VETO Unit:** collects information from VETO sources and processes this data in order to classify each bunch for each individual output (e.g. VETO User). Classification of bunches can be based in algorithms outputs, implemented in the VETO unit, like threshold detection, comparison and evaluation of logical connections. Possible classes could be:
 - **VETO:** measurement of this bunch was most likely not successful
 - **NO-VETO:** neither as VETO nor as GOLDEN



- **GOLDEN:** bunch measurement was very good
- **VETO User:** based on the VETO Unit decision, the VETO User device (like fast 2D imaging detectors as presented and discussed in the report on online 2D image processing) will reject, remove or tag the data set for the related bunch and takes further required steps depending on the implementation of the individual device and its DAQ chain.

The information from the VETO system has to be delivered with as low latency as possible, as this it has to be further processed and transmitted to the front-end electronic of the VETO user and applied to the storage cell as fast as possible in order to make efficient use of it. A Low Latency Link (LLL) protocol was develop to be the high-speed link with which the subsystems of the European XFEL can distribute the acquired beam information and react on it at the receiver side with minimal latency. Implementation is based in Xilinx Logicore GTP core, with parameters being set to achieve low latency.

VETO Unit Prototype

A VETO Unit prototype has been implemented in a DAMC2 Board, a compact and economical solution aimed at several applications for control and Data Acquisition systems at the European XFEL accelerator. The board includes four SFP connectors, for external High-speed optical interfaces.

The current implementation allows for four VETO sources to connect to the VETO Unit. Per VETO source, data limits can be configure to classify each word in a data packet as Gold, VETO or NO-VETO. A decision is taken on a data packet, which is a combination of each individual data word classification (up to five words). Per data packet, all VETO sources decisions can (but don't have to) be combined and use as an input for a look up table. This allows VETO decisions to be based on information from multiple VETO sources. The functionality of the VETO unit was successfully tested in hardware.

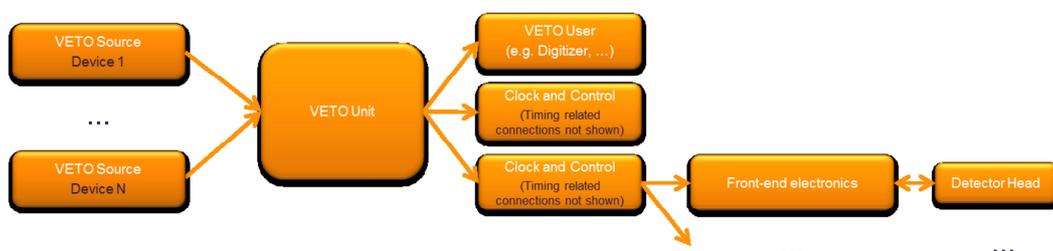
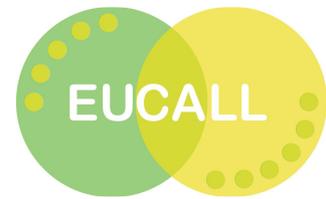


Figure 7: Overview of VETO system



5.3 Data reduction

For example, European XFEL runs with a 10 Hz trigger signal which is synchronized for all part of facility. This trigger signal is followed by 2700 pulse with the frequency of 4.5 MHz. Electrical pulses which are generated by detectors like PES system or photo diodes are very short pulses with pulse width in the order of several ns. These pulses can be also very sharp in rising edge.

To be able to process these pulses, very fast digitizers as described above are required. Digitizing these pulses lead to a very high data rate which should be transferred to software to process and analyze x-ray properties of the light. Transferring this amount of data takes a lot of time and also need huge amount of resources like memory.

To avoid transferring raw data directly to software different data reduction algorithms can be implemented in FPGA directly after digitizing to reduce the amount of data. When raw data form ADC is processed in real time inside FPGA, information will be transferred to software for more detail analysis.

Three different data reduction algorithms are implemented in FPGA to reduce amount of data. These algorithms are 1- zero suppression 2- Peak detection 3- Energy calculation. Each of them is described in detail in the following section.

6. Implemented processing algorithms

6.1 Energy detection

This algorithm is implemented to measure pulse integration which will be used to analyze the properties of the light (or particle) pulse.

As an example, consider four APD detectors used by FXE experiment at European XFEL. These detectors are positioned two in X and the other two in Y direction. X-ray beam passes through the Diamond screen which scatters light and the photons are detected by these APD detectors. APD detectors generate electrical pulses for each detected pulse. These pulses are sampled by digitizers and processed for monitoring pulse intensity and position.



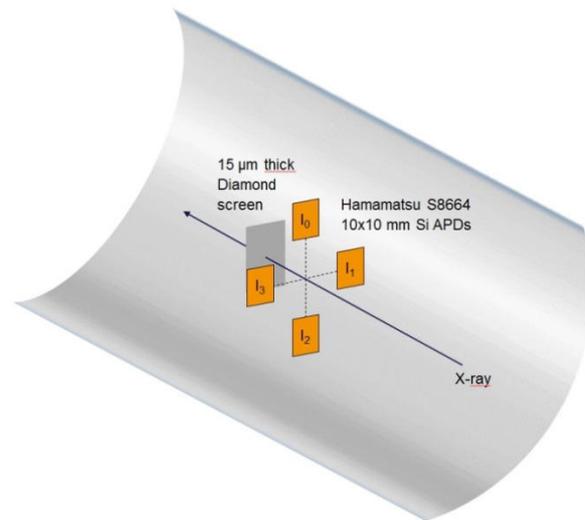


Figure 8 - Position of APD detectors in the FXE Experiment (European XFEL) for calculation of beam intensity and position

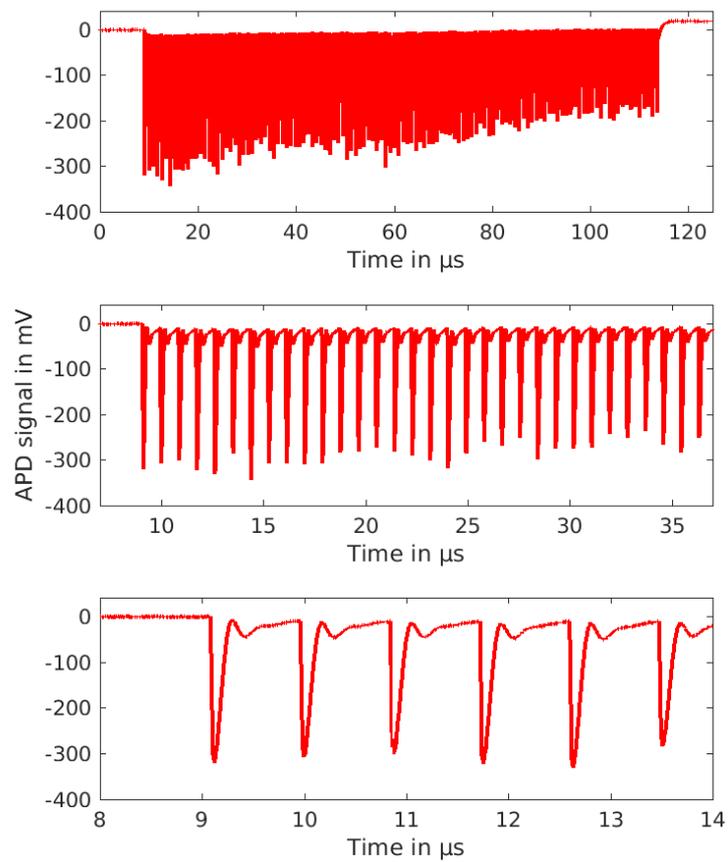


Figure 9: Plots of the digitized signal in different zoom factors.

In Figure 9, detected pulses by digitizers are shown. By processing the detected pulses from detectors, it will be possible to know intensity and the position of each beam as Pulse energy $\sim \sum_i I_i$, x-position $\sim I_3 - I_1$ and y-position $\sim I_0 - I_2$.

In this type of experiment, because pulses are generated with a pre-defined frequency and also are synchronized with trigger (see Section 3.1). This makes implementation of integration algorithm easier in the FPGA: since the position of all detected pulses by detectors respecting to the trigger is known in advance, based on a defined area the hardware can start and stop integration based on that value. Pre-analyze of raw data out of ADC in digitizers is necessary to extract the area based on sample number. Data from the ADCs are also synchronized with the trigger and the samples are indexed starting from zero for the first sample referenced to trigger. By receiving the raw data and analyzing the first pulse, it will possible to define the pulse area according to sample number. In addition to be able to get the area of next pulse it is only enough to know the frequency of the pulses. The last parameter which can be extracted from raw data is called initial delay, the time between the trigger occurrence and first pulse. This information will be used by hardware to calculate the integration for all pulses in each train. All these parameters are shown in Figure 10, where the first pulse of a train is depicted.

There are two different type of Analog to Digital Converters in the terms of their front end circuit: AC-Coupled and DC-Coupled. AC-Coupled front ends use transformers to bias the ADC input; on the other hand DC-Coupled front ends deploy differential amplifiers for coupling ADCs. Each of these techniques has their own advantages and disadvantages. The main difference between them is that in AC-Coupled digitizers only AC component of input will be transferred, while in DC-Coupled ones both AC and DC components of input signal will be transferred to ADC to digitize. This reality influences how the integration calculation is perform in hardware.

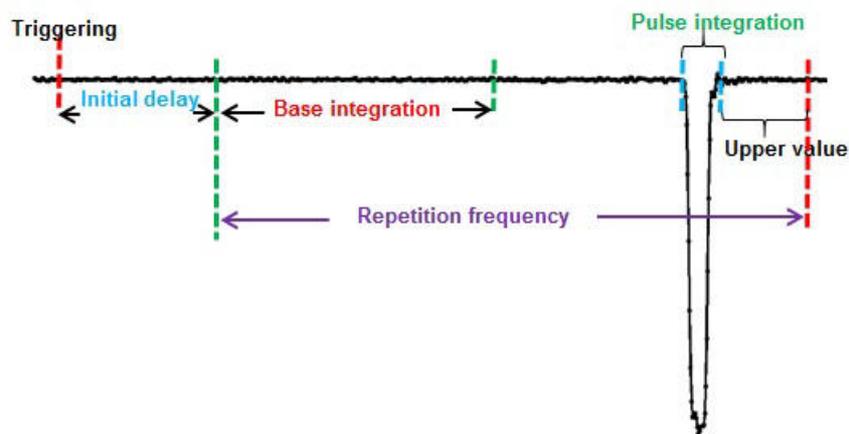
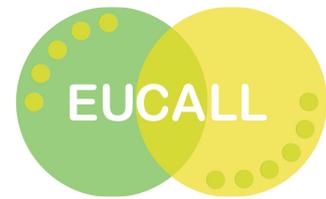


Figure 10: Parameters for the Energy detection algorithm



Not having the DC component of input signal in AC-Coupled digitizers plays an important role especially specially when the input is a burst of pulses in a limited time window and having DC offset. When there is no pulse, the front end output (the ADC input) will be in a steady state of noise level, also known as base line. As soon as the burst of pulses arrive the system will try balance it and remove the DC component, which will appear as a drift from the base line in the beginning. This drift in the beginning will result to a wrong value of pulse energy for first pulses. On the other side, the DC-Coupled digitizers sample both DC and AC components of input signal. If the input signal has DC offset, the output of integration algorithm will have wrong value for all pulses.

To be able to support both AC-Coupled and DC-Coupled digitizers, integration of noise level is necessary. To get precise pulse integration values, the sum of the noise level samples should be removed from the pulse integration. A noise area can be defined for each individual pulse, which is will be the noise area (time window) before the pulse. The wider noise area, the better estimation of noise energy is possible. If the noise area is not equal to the pulse area, its integration should be scaled to be correctly removed from the pulse integration.

When all parameters are provided to hardware, integration will be done in real-time. Scaling of noise integration is selectable to be done either in hardware or in software. The result of integration (noise area and pulse area) is sent to the software for more processing.

In the experiments that the position of the pulses is not known in advance, peak detection algorithm should be used. In this case, the pulses should be detected and then further processing can be done.

6.2 Peak detection

Peak detection algorithm is implemented next to other algorithms inside FPGA and runs in parallel to other algorithms. The result of this algorithm is also used for analyzing x-ray properties of light. Data source to all algorithms is the same synchronized to external trigger. This algorithm detects pulses in a train for each trigger and then looks for maximum value in the pulse. When the maximum value is detected, it is sent along with several other parameters to software.

This algorithm is implemented to work on parallel input, like SP-Device ADQ412 digitizer which has four channels with eight samples per clock cycle per channel. It processes eight samples per clock cycle and detects maximum two peak values out of these eight samples.

In addition of actual eight samples per clock cycle, this algorithm needs four ancillary samples, time wisely two samples on the right side and two other samples on the left side. These ancillary samples which are on the left side are two older samples than actual samples and two on the right side are newer adjacent samples. To find the peak area first differences between every other sample are calculated then these differences are compared to zero to



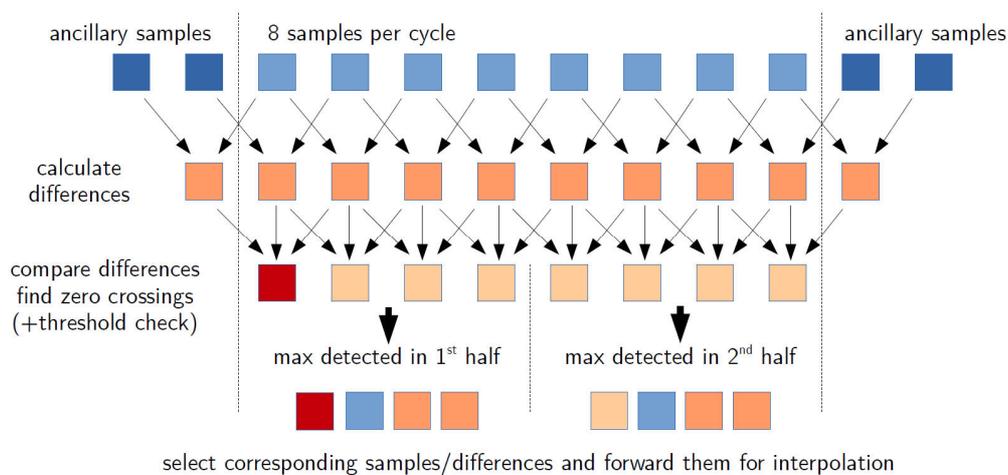


Figure 11: Peak detection dataflow

find zero crossing. This result is valid only when the samples are above threshold. Predefined threshold value is a value above noise level which is defined via software after analyzing raw data and finding noise level.

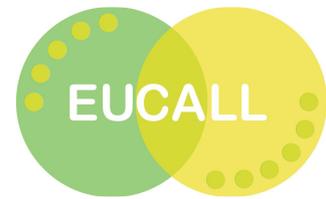
In Figure 11, eight actual samples are presented in blue color with four ancillary samples, two on the right side and two on the left side with dark blue color. First step is getting difference between every other sample which leads to ten values. These ten values are analyzed to find zero crossing. The analyze outcome will be eight values.

These eight values will be checked to find peak value. As mentioned above for each eight samples only two peak values are detectable. Therefore for each detected peak four samples are reported and peak value will be the second sample from the right out of these four samples along with index of the peak respect to the trigger signal.

6.3 Zero suppression

This algorithm is implemented in FPGA along with other algorithms as data reduction algorithm. Actually this algorithm does not do any data processing but only data reduction. Output of this algorithm is used for debugging purpose. This algorithm suppresses noise information and as soon as a pulse is detected collects pulse area samples and send these samples to software with pulse timing information respecting to trigger signal.

The method which is used in this algorithm is similar to peak detection with the difference that in peak detection, when rising edge is detected, looking for peak value is started, but in zero suppression as soon as rising edge is detected collecting data will be started. Collecting data should be limited to some value, which can be done in two ways. First way is to detect falling edge of pulse and stop collection with falling edge; second way is to define a window via software to dictate how many samples should be collected after rising edge. The difference between these two methods is the way of data transfer to software. In first way packet length can be different pulse to pulse depends on pulse width, in the second way



packet length will be fixed because the size of window is fixed and defined by software in the beginning.

Another parameter which is used in this algorithm is threshold value. All samples are compared to this value which is a little bit above to noise level. When several samples are above threshold value, pulse detection will be reported, and then sample collection will be started.

In this algorithm defining threshold is done in two different ways, first way is similar to peak detection algorithm, defining threshold value via software. In this way at first raw data from ADC should be analysed and noise level should be detected then a value can be introduced as threshold value. Second method is using adaptive threshold detection. This method is also implemented in Hardware and processes all data in real time and detects noise level. This detected noise level is incremented with a value to be used as threshold value.

6.4 Pile-up compensation

In this section we describe the development and implementation in FPGA of an algorithm for pile-up event treatment, together with its limits and possible future improvements. It is applied for processing the signals produced by gamma ray interactions in LaBr₃:Ce scintillator detectors. However, as presented below, the algorithm is very general and can be applied to signals from many types of detectors. For isolated events, the algorithm output corresponds (strictly) to a digital QCD (charge to digital convertor performing signal integration) as implemented in various commercially available firmware distributed with digitizers. The new algorithm is enhancing this common method:

- (i) detecting when two interactions very close in time are generating the pile-up event
- (ii) extending the integration window such to provide correctly the total charge corresponding to both interactions
- (iii) estimating the charge associated the first event.

The algorithm is based on the stability of detector pulse shape: meaning that signals are similar (up to same fluctuations due to noise) if the energy deposited in detector is the same. Actually, it can be check in practice that, up to a multiplying factor, the signals are similar over a wide range of deposited energies in the detectors.

Modeling of scintillator signals with simple functions such as bi-exponential convoluted with a Gaussian was studied in [1]., where the two exponential are due to rising and decay time constants of the scintillation mechanism while Gaussian convolution accounts mainly for photomultiplier (PMT) resolution. The feasibility of applying a digital pulse deconvolution for bi-exponential signal shape (neglecting the Gaussian convolution) was further studied [2] for NaI scintillators. By moving as much of the pulse processing in FPGA, this method has the



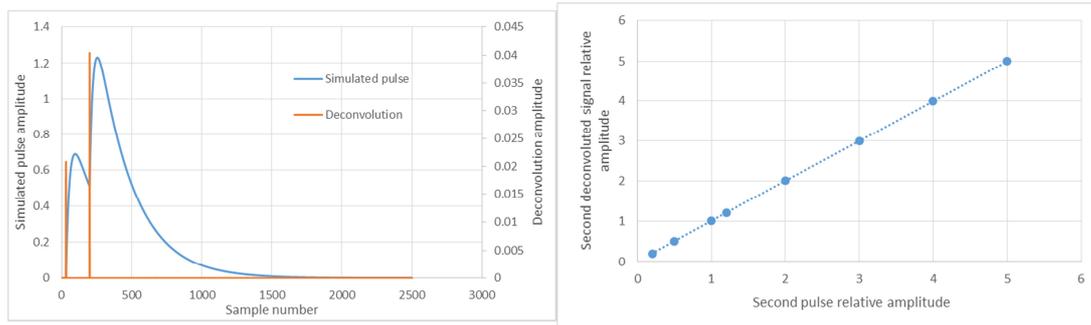


Figure 12: Left: Deconvolution of a noise-free simulated pulse. Right: Linearity of the deconvolution algorithm for noise-free pile-up pulses.

advantage that it significantly decreases the data throughput necessary between the digital front end and the storage system which is a limiting factor for very high speed data acquisition system. This method promises the additional advantage that the pulse produced by a scintillator and PMT can be deconvoluted into a very narrow delta-like signal and as such it treats pulse pile-up by removing the slow-decaying tail of the pulse, in theory improving the maximum event rate significantly. Figure 12 (left) shows a series of two biexponential pulses, which can approximate the scintillator and PMT signal and the deconvolution result. The amplitude ratio of the pulses is 1:2, and this is the same ratio that the integrals of the two deconvoluted peaks produce.

One can check that the linearity of this procedure is preserved by varying the relative amplitudes of the piled-up pulses and Figure 12 (right) shows that in the case of completely noise-free signals this linearity is perfectly preserved.

Unfortunately, adding noise to the original pulses starts degrading the output signal to the point it becomes undistinguishable from the background. Figure shows the results, making pulse parameter recovery impossible just from this signal.

Applying a low-pass filter to this signal yields significantly better results, as most of the noise introduced by the deconvolution step is in the high frequency domain. Figure 14 shows the filtered result of the very noisy signal in Figure 13 whereas Figure 14 focuses on a higher noise amplitude, emphasizing the performance of the filter and linearity of the whole system in this case.

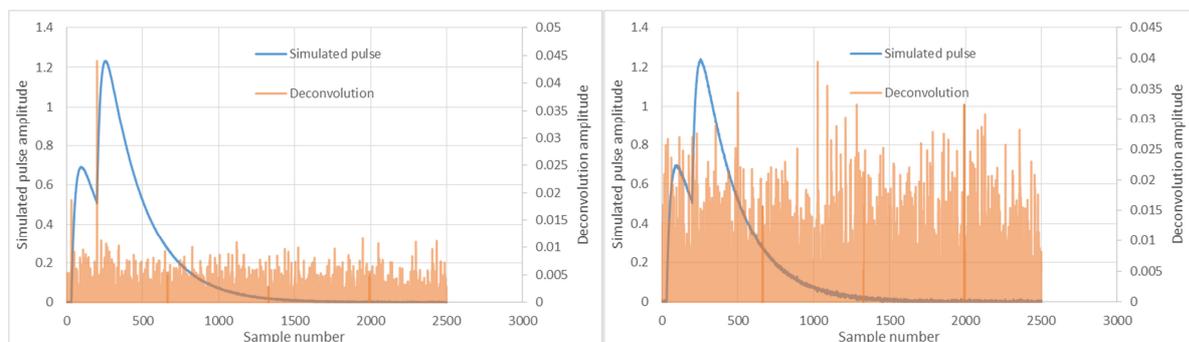


Figure 13: Deconvolution of noisy simulated pulses, with different Gaussian noise amplitudes, standard deviation 15% of pulse amplitude. Left: amplitude 1% of the pulse amplitude. Right: amplitude 3% of the pulse amplitude.

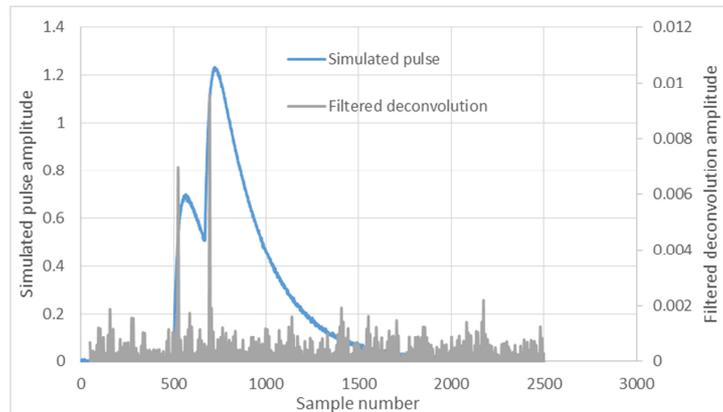


Figure 14: Filtered deconvoluted signal, with noise amplitude 3% of pulse amplitude, same conditions as those presented in Figure 13 (right).

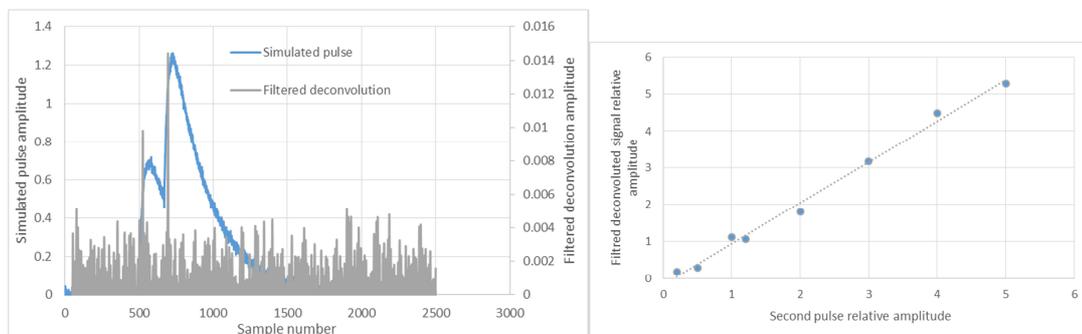


Figure 15: Filtered deconvoluted signal with noise amplitude 10% of pulse amplitude and linearity of the filtered signal in this case.

The pulse deconvolution step together with the finite impulse response (FIR) filter of order 50 that has been used for these tests has been implemented in Verilog and synthesized for a Kintex7 FPGA. These two steps can be run up to 350 MHz according to synthesis. Results obtained from the behavioral model simulation using Xilinx tools are presented in Figure 16, where the filtered deconvolution output suffers from some quantization errors as the computation steps are implemented using integer arithmetic.

The above method applied successfully in [2] for NaI scintillators is not expected to work equally well for LaBr₃:Ce detectors that are one order of magnitude faster and the Gaussian convolution applied on the biexponential cannot be neglected in pulse shape modeling. Therefore, an algorithm for arbitrary signal shape was studied and implemented.

The shape of LaBr₃:Ce scintillator signals and its stability was measured as described in the following. A system comprising of a 2" x 2" cylindrical LaBr₃:Ce crystal coupled to a photomultiplier tube (PMT) was used in test measurements with a ⁶⁰Co calibration gamma source. Gamma rays with energies 1173 keV and 1332 keV are associated with the ⁶⁰Co gamma source decay radiation [3] and they interact with the crystal via photoelectric effect, Compton scattering and pair production. The processes result in signals with different total

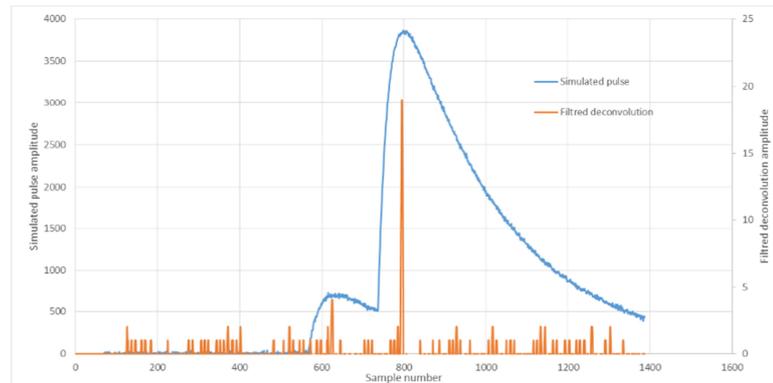


Figure 16: Simulation of filtered deconvolution output using the FPGA code showing some quantization error.

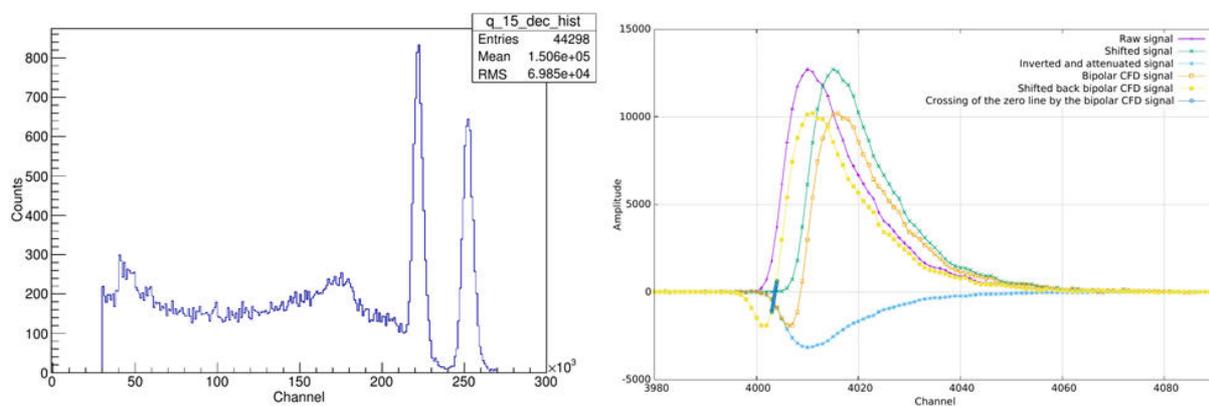


Figure 17: Left: A ^{60}Co gamma ray source energy distribution based on integration of the total area of the recorded signals. Right: Application of the CFD method to determine the interaction time.

area and amplitudes and which were processed and recorded using a 500 MS/s SP Devices ADQ14DC-2A digitizer.

The amplitude of the recorded signals and their total integrated area are related to the energy deposited in the crystal. Hence, an energy distribution can be constructed on the basis of one of these parameters. Such a distribution is presented in Figure 17 – the histogram represents the number of recorded signals with a given total integrated area. A lower limit for the area was set and it refers to the starting point of the distribution. The shape of the distribution clearly shows the full absorption peaks related to the 1173 keV and 1332 keV gamma rays and the Compton continuum. The good separation of the two peaks, corresponding to a resolution of about 3% FWHM, is due to very high scintillation yield of $\text{LaBr}_3:\text{Ce}$ as compared to other scintillators types.

The set of recorded data was used to determine a reference $\text{LaBr}_3:\text{Ce}$ signal and study the variations of individual signals from its shape. The procedure requires normalization with respect to the total integrated signal area and a precise alignment of the moment of detection of each of the incoming pulses. Thus, for each of the recorded signals a procedure based on the Constant Fraction Triggering (CFD) method [4] was performed in order to determine the time of detection.

The application of the CFD method for a sample pulse is presented in right graph in Figure 17. The raw LaBr₃:Ce signal is used to construct two other pulses – one of them is identical to the initial one but delayed by a value of t_{cfd} , where t_{cfd} is the rise time of the signal from a given fraction k of V_{max} (its maximum value) to V_{max} ; the other signal is obtained by inversion and attenuation of the initial one by a factor equal to k . The sum of these signals result in a bipolar pulse with a zero-crossing point which determines the moment when the shifted pulse reaches the value of kV_{max} .

Values of $k = 0.25$ and $t_{cfd} = 10$ ns (5 channels) were used with the present dataset of LaBr₃:Ce signals. The total integrated area of each of the pulses was determined in the time range between $t_{max} - 6t_{cfd}$ and $t_{max} + 15\tau$, where $\tau = 16$ ns indicates the decay time of LaBr₃:Ce. The zero-crossing of the bipolar CFD pulse was determined with the precision of 1/16 of a channel. Thus, all recorded signals were aligned in time to a value with such an uncertainty.

The time aligned and total area normalized pulses were also normalized with respect to the position of the 1/16 fraction of the channel of their CFD bipolar signal zero-crossing. An average LaBr₃:Ce signal was determined after summation of all normalized signals and a division by the total number of events. The obtained average LaBr₃:Ce signal was used as a reference signal in the algorithm presented further below. All the time aligned and area normalized pulses were compared to it. The variations for each channel with respect to the reference signal were calculated and summed. This allowed determining an uncertainty of the value for each channel. The average signal and the respective variations for each channel are presented in Figure 18. One can observe that the variations for each channel are small and, respectively, the shape of the signal is stable.

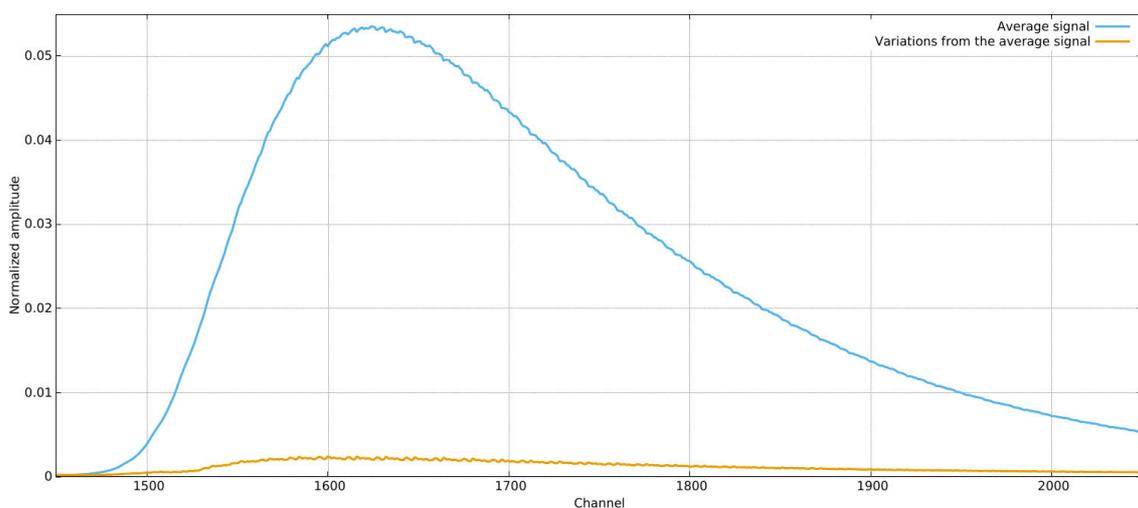


Figure 18: A reference LaBr₃:Ce signal and its uncertainty in each channel. It was determined by averaging a set of time aligned and total area normalized LaBr₃:Ce pulses.

The stability of the shape of the $\text{LaBr}_3\text{:Ce}$ signals with respect to the deposited energy in the scintillator is the second condition for applicability of developed algorithm. In order to investigate it, an average signal was determined in a similar way, with a restrictive condition applied for the total integrated areas of the signals. The condition was chosen in such a way that only pulses related to the full absorption peaks at 1173 keV and 1332 keV are considered while the events from the Compton background are dismissed. A comparison between the average signals obtained with and without the area condition is presented in Figure 19. The almost perfect overlap of the two average pulses confirms that the reference $\text{LaBr}_3\text{:Ce}$ signal is stable with respect to energy of the incoming gamma rays within a broad energy interval.

Starting from the reference signal: $\{\chi_i \geq 0, i = 1, \dots, L\}$ normalized to unity:

$$\sum_{i=0}^L \chi_i = 1$$

one defines a factor related to partial integral of the reference signal:

$$R[j] = \sum_{i=0}^j \chi_i$$

such that, when the partial integral of a real signal $\{s_i, i = 1, \dots, L\}$ is:

$$p[j] = \sum_{i=0}^j s_i$$

the total signal integral (corresponding to energy deposited) is predicted to be:

$$E = \frac{p[j]}{R[j]}$$

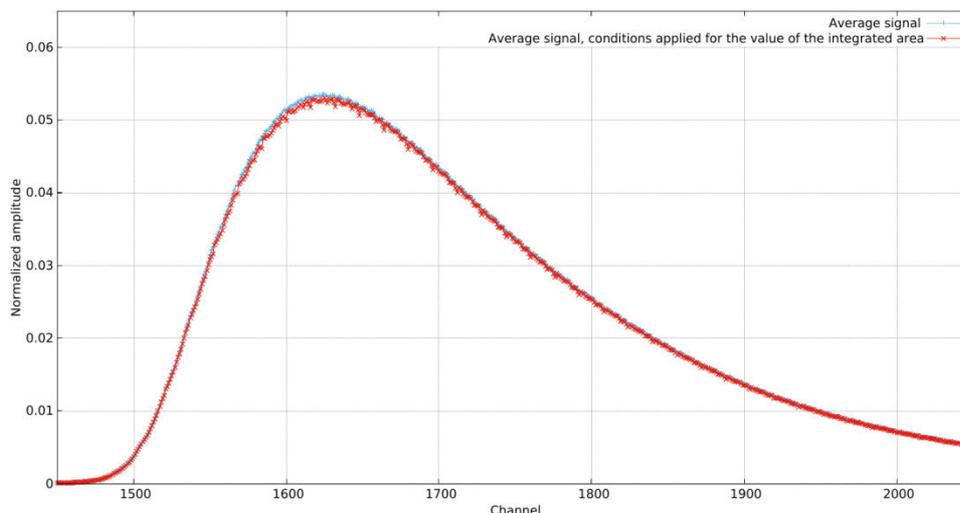
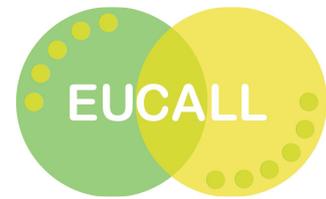


Figure 19: A comparison between average $\text{LaBr}_3\text{:Ce}$ signals, obtained with and without a condition set for the total integrated area. The overlap of the signals suggest a stability of the shape of the $\text{LaBr}_3\text{:Ce}$ pulse with respect to the energy deposited in the scintillator



Thus, the energy of the first event can be obtain using only the first part of signal, till the second overlapping event starts and alter the signal due to first event. Larger delay between the signals means that larger part of the first signal is used in energy calculation such that $1/R[j]$ correction is almost 1 and the resolution of calculated energy closer to optimum. In order to detect the occurrence of the second signal, the deviation of the real signal from reference signal can be evaluated at each new sample:

$$d_j = s_j - \chi_j \frac{p[j - 1]}{R[j - 1]}$$

and a threshold condition applied. When signal overlap is detected, the current estimation of first signal energy is saved and integration continued for more L samples such that the total integral is calculated, that is the sum of the two deposited energies

As described, the algorithm is applicable to any type of reference signal. It is also easy to generalize for more than two signal overlap with an additional degradation of resolution that has to be evaluated case by case. The algorithm was checked using spreadsheets and simulated with pulses prepared by adding two references pulses multiplied with two different values and delayed, plus scalable level of noise generated according to Poisson distribution.

The algorithm was developed and implemented for Kintex-7 FPGA open to users in the SP Devices digitizer mention above. It was coded using Verilog language modifying one the two modules available for user firmware customization as part of development kit provided by SP Devices together with digitizers. The arrival of the first signal is detected by level trigger already implemented in the firmware, and then calculated with 1/16 sampling period accuracy using the same digital CFD algorithm (and parameters) as used for reference signal measurement. The two arrays of constants $1/R[j]$ and $\chi_j/R[j - 1]$ are loaded in FPGA memory (using `$readmemb()` Verilog instruction) with a length of about 2000 because 16 values are needed for each sample. The 32 bit floating point representation with 24 bits for mantissa was used for these constants, thus matching the 18 x 25 bit multiplication available in DSP cells. The total amount of memory needed for the two arrays (128 kb) is not a problem considering the large on chip memory of Kintex-7.

The real implementation is further complicated because the FPGA is run at 250 MHz meaning that two samples are processed in parallel. The behavioral simulation used same input waveforms as the spreadsheet allowing checking in detail the synchronization of values and applying correct buffering in order to obtain the same output values.



7. Usage of developed solutions

7.1 Measurements at PETRA III and European XFEL



Figure 20: Top view of PETRA III synchrotron

PETRA III is the high brilliance third Generation Synchrotron Radiation Source at DESY, located in Hamburg, Germany. With the circumference of 2.3 km PETRA III is the biggest and most brilliant storage ring light based x-ray source for high energy photons providing a brilliance exceeding 10^{21} ph/(s mm² mrad² 0.1% BW).

The beamlines at PETRA III are distributed over three experimental halls: ‘Max von Laue’ which is the largest with 300 meter long experimental hall and covers one octant of the 2304 meter long PETRA storage ring on DESY site, the ‘Paul P. Ewald’ and the ‘Ada Yonath’ halls, located on the northern and eastern side of the experimental hall ‘Max von Laue’. On the 7000 m² large experimental floor of ‘Max von Laue’, 15 beamlines are operated by DESY, Helmholtz-Zentrum Geesthacht (HZG), and the European Molecular Biology Laboratory (EMBL) with more than 30 experimental stations which have been optimized for the use of the high brilliance of the PETRA III beam.

P01 beamline specializes in Nuclear Resonant Scattering (NRS) and Inelastic x-ray scattering (IXS and RIXS) experiments at photon energies between 2.5 keV and 80 keV. The beamline offers high energy resolution in the range from 1 meV to about 1 eV and high spatial resolution in the (sub-) micron regime. At P01 beamline, pump-probe experiments with pulse limited temporal resolution (100 ps) was perform using SP-Devices ADQ412 digitizers with an early version of Energy Detection algorithm (see Section 6.1).

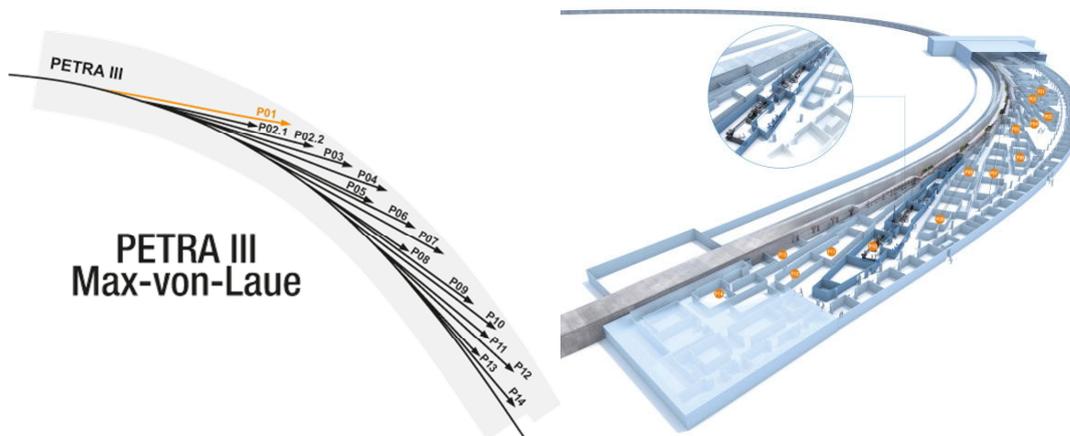


Figure 21: PETRA III beamlines (Max-von-Laue Hall)

In the experiment molecular systems in liquid solutions are excited by the pulsed laser source and the total x-ray fluorescence yield (TFY) from the sample is recorded using silicon avalanche photodiode detectors (APDs). The experiment was done in 60 bunch mode by exploiting a synchronized 3.9 MHz laser excitation source.

The subsequent digitizer card samples the APD signal traces in 0.5 ns steps with 12-bit resolution. These traces are then processed to deliver an integrated value for each recorded single x-ray pulse intensity and sorted into bins according to whether the laser excited the sample or not. For each subgroup the recorded single-shot values are averaged over 107 pulses to deliver a mean TFY value with its standard error for each data point, e.g. at a given x-ray probe energy [5]. The successful results taken from these tests at PETRA III led to the redesign and improvement of the Energy Detection algorithm to be able to support high rate pulses.

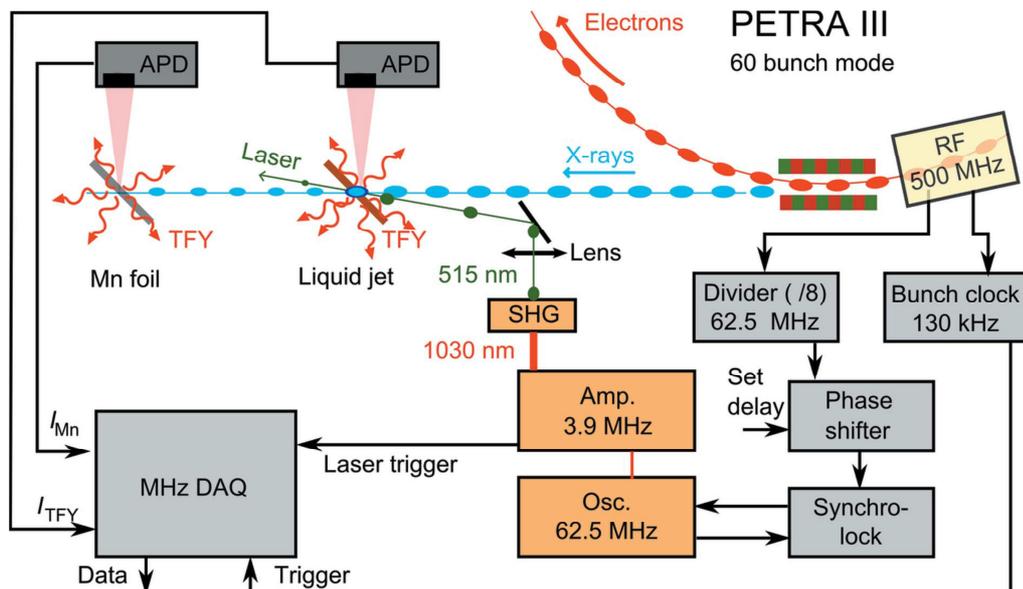


Figure 22: Sketch of the experimental setup at the dynamics beamline P01 of PETRA III

The Energy Detection algorithm is currently used by FXE experiments at European XFEL. As described in Section 6.1, the setup includes four APD detectors and burst of 120 pulses at 1.1 MHz. The output of this algorithm is used for Intensity and beam Position Monitoring.

The Energy Detection and the Peak detection (see Section 0) algorithm are used in the PES system at European XFEL. The PES is a diagnostics device whose purpose is to provide both machine operators and users at the experimental stations about light pulse properties. It has 16 electron time-of-flight spectrometers placed in a circle orthogonal to the x-ray laser (see Figure 23). In the center a gas (N_2 , Argon, Krypton or Xenon) is injected, which is ionized by the x-ray laser. When an atom or molecule is ionized, photo-electrons are emitted in all directions with an angular distribution that depends on the gas and the photon energy.

Photon-energy is related to the electrons kinetic energy: the higher the energy, the shorter will be the time for the electron to reach the detector. From the temporal spectrum, the photon energy will be known. Peak detection algorithm is used to calculate the photon energy. To determine the polarization of beam the following approach is used:

- If photon polarization is linear the angular distribution will be very anisotropic.
- If photon polarization is circular the angular distribution will be very isotropic.

By measuring the intensity ratio on all 16 detectors, the polarization of light can be determined. For this purpose, the Energy Detection algorithm is used.

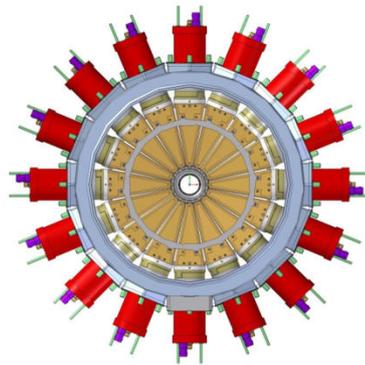


Figure 23: PES system

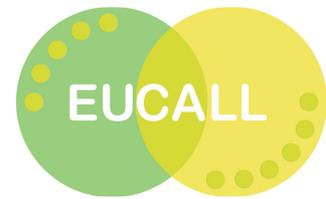
7.2 Measurements from ELI-NP

The performance of implemented algorithm was measured with real pulses in a set-up consisting on two detectors connected, respectively, to the Channel A and Channel B input of the digitizer. One of the pulses can be easily delayed with up to few tens of nanosecond using cable of various lengths, while the overlap is done by simple sample addition in the FPGA in the beginning of implemented algorithm. In order to avoid the development of an application to readout the modified FPGA firmware, a simple solution has been applied using the ADCaptureLab application provided by the vendor and altering the output waveform for the Channel B. The Channel A was used to output the waveform resulted from input signal summation

8. Conclusion

FPGA platforms allows for multiple algorithms to be implemented at the front-end of detectors, allowing for data processing and analysis to be perform at acquisition time. From the descriptions presented, it is clear that FPGA programming is time consuming and requires a deep understanding of not only of the platform but also of the physics of the signal to be process as well as the algorithm to be implemented.

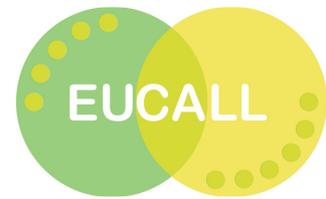
Different algorithms have been implemented which are currently in use in the different facilities, providing results for diagnostics, beam alignment and position, data rejection and pulse characterization. Since this data analysis is available at a very early stage, this allows for a more efficient use of beam time, optimization of laser performance or stability and reduction of data storage.



References and publications

- [1] W. Xiao, A.T. Farsoni, H. Yang, D.H. Hamby, *Model-based pulse deconvolution method for NaI(Tl) detectors*, Nuclear Instruments and Methods in Physics Research Section A, Volume 769, Pages 5-8 (2015); doi: <https://doi.org/10.1016/j.nima.2014.06.022>
<https://doi.org/10.1016/j.nima.2014.09.034>
- [2] W. Xiao, A.T. Farsoni, H. Yang, D.H. Hamby, *A new pulse model for NaI(Tl) detection systems*, Nuclear Instruments and Methods in Physics Research Section A, Volume 763, Pages 170-173 (2014); doi: <https://doi.org/10.1016/j.nima.2014.06.022>
- [3] National Nuclear Data Center (<http://www.nndc.bnl.gov/>)
- [4] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments*, ISBN 3-540-17386-2 Springer-Verlag Berlin Heidelberg New York (1987).
- [5] A. Britz, T.A. Assefa, A. Galler, W. Gawelda, M. Diez, P. Zalden, D. Khakhulin, B. Fernandes, P. Gessler, H. Sotoudi Namin, A. Beckmann, M. Harder, H. Yavaş, C. Bressler, *A multi-MHz single-shot data acquisition scheme with high dynamic range: pump-probe x-ray experiments at synchrotrons*, J. Synchrotron Rad, 23 1409–1423 (2016); doi: <http://dx.doi.org/10.1107/S1600577516012625>
- [6] Gabor Gyulai: Master Thesis – Distributed RF over White Rabbit Network, BME 19.01.2018
- [7] Maciej Lipinski: White Rabbit – Ethernet-based solution for sub-ns synchronization and deterministic, reliable data delivery, http://www.ieee802.org/802_tutorials/2013-07/WR_Tutorial_IEEE.pdf
- [8] Seven Solutions: WR-ZEN TimeProvider – User Guide (15 April 2016)
- [9] Open Hardware Repository: Simple PCIe FMC carrier (SPEC) <https://www.ohwr.org/projects/spec/wiki/wiki> (November 2017)
- [10] Javier D. Garcia-Lasheras: Getting Started with the SPEC – A tutorial for the Simple PCI Express Carrier project newcomers (Version 1.0, March 2014, CERN BE-CO-HT)
- [11] Open Hardware Repository : FMC cards developed as Open Hardware, http://www.ohwr.org/projects/fmc-projects/wiki/OHR_developments (May 2017)
- [12] Tomasz Wlostowski: Fine Delay Design Notes (June 2013), CERN BE-CO-HT
- [13] Alessandro Rubini and Federico Vaga: ZIO User Manual (version-1.0) (January 2013) CERN BE-CO-HT
- [14] Tomasz Wlostowski, Alessandro Rubini: Fine Delay User’s Manual (April 2014), CERN BE-CO-HT
- [15] Open Hardware Repository : FMC DEL 1ns 4cha – Software, http://www.ohwr.org/projects/fine-delay-sw/wiki/Release_v2014_04 (May 2017)
- [16] Open Hardware Repository: fmc-bus, <https://www.ohwr.org/projects/fmc-bus> (November 2017)
- [17] Open Hardware Repository: White Rabbit Network Interface Card, <https://www.ohwr.org/projects/wr-nic/wiki> (November 2017)





- [18] Javier Díaz & Miguel Jiménez (Univ. of Granada), Rafael Rodriguez & Benoit Rat (Seven Solutions): White-Rabbit NIC Gateway (17 February 2014 - wr-nic-v2.0)
- [19] Open Hardware Repository: WR Streamers <https://www.ohwr.org/projects/wr-cores/wiki/wr-streamers> (2017 November)
- [20] Tomasz Wlostowski: White Rabbit Trigger Distribution a.k.a. MIMO systems using WR (ICALEPCS 2017)
- [21] Open Hardware Repository: LHC Instability Trigger Distribution project (LIST) https://www.ohwr.org/projects/wr-node-core/wiki/LHC_Instability_Trigger (November 2017)
- [22] Open Hardware Repository: Distributed Direct Digital synthesis over White Rabbit <https://www.ohwr.org/projects/wr-d3s/wiki> (November 2017)
- [23] Tomasz Wlostowski: Distribution of RF signals using WR (ICALEPCS 2017)
- [24] Open Hardware Repository: FMC DAC 600M 12b 1cha DDS <https://www.ohwr.org/projects/fmc-dac-600m-12b-1cha-dds/wiki> (November 2017)

